Respiratory Medicine (2011) 105, 1507–1515

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/rmed

Quality of Spirometry tests performed by 9893 adults in 14 countries: The BOLD Study

P. Enright [a], W.M. Vollmer [b], B. Lamprecht [c,*], R. Jensen [d], A. Jithoo [e], W. Tan [f], M. Studnicka [g], P. Burney [h], S. Gillespie [b], A.S. Buist [i]

[a] The University of Arizona, Tucson, AZ, USA
[b] Kaiser Permanente Center for Health Research, Portland, OR, USA
[c] Paracelsus Medical University, University Hospital Salzburg, Department of Pulmonary and Critical Care Medicine, Muellner Hauptstrasse 48, 5020 Salzburg, Austria
[d] Latter Day Saints Hospital, Salt Lake City, UT, USA
[e] National Heart and Lung Institute, Imperial College, London, UK
[f] iCapture Center for Cardiovascular and Pulmonary Research, University of British Columbia, Vancouver, BC, Canada
[g] Paracelsus Medical University, Department of Pulmonary Medicine, Salzburg, Austria
[h] National Heart and Lung Institute, Imperial College, London, UK
[i] Oregon Health and Sciences University, Portland, OR, USA

Received 13 February 2011; accepted 12 April 2011
Available online 6 May 2011

KEYWORDS
Spirometry;
Quality control;
COPD

Summary
Objective: to determine the ability of participants in the Burden of Obstructive Lung Disease (BOLD) study to meet quality goals for spirometry test session quality and to assess factors contributing to good quality.
Methods: Following 2 days of centralized training, spirometry was performed pre- and post-bronchodilator (BD) at 14 international sites, in random population-based samples of persons aged ≥40 years, following a standardized protocol. The quality of each test session was evaluated by the spirometer software and an expert reading center. Descriptive statistics were calculated for key maneuver acceptability variables. A logistic regression model identified the predictors of acceptable quality test sessions.
Results: About 96% of test sessions met our quality goals for a low back-extrapolated volume (BEV), time to peak flow (PEFT), and end-of-test volume (EOTV). The mean forced expiratory time (FET) was 10.4 s. Ninety percent of the maneuvers with the highest FVC had a forced expiratory time (FET) > 6.8 s. About 90% of test sessions had FEV_1 and FVC which were repeatable within 150 mL. Test quality was slightly better for post-BD test sessions when compared to pre-BD. Independent predictors of adequate test quality included female sex, younger age, higher education, lack of dyspnea, higher pre-BD FEV_1, less BD responsiveness, and study site.

* Corresponding author. Tel.: +43 662 4482 3300; fax: +43 662 4482 3303.
E-mail address: b.lamprecht@salk.at (B. Lamprecht).

doi:10.1016/j.rmed.2011.04.008

Conclusions: Quality goals for spirometry tests were met about 90% of the time in these population-based samples of adults from several countries.

Introduction

The airway obstruction of COPD is determined by a low post-bronchodilator (post-BD) FEV_1/FVC and the severity of the airway obstruction is determined by the FEV_1 (percent predicted). Poor quality spirometry tests can cause either a falsely high or a falsely low FEV_1/FVC (false negative or false positive interpretation of airway obstruction) and an under-reported FEV_1 (more severe airway obstruction).

The current ATS/ERS goals for acceptable and repeatable spirometry tests[1] are based on the ability of well-trained technologists to meet these goals in 9 of every 10 adult patients referred for testing in a single hospital-based pulmonary function laboratory in the United States.[2] Spirometry quality has been reported from a large study of workers participating in the World Trade Center Responders program in New York City.[3] However, very little has been published about the ability of population-based samples of adults from many countries to meet these goals for spirometry done post-bronchodilator (as for COPD case-finding). The storage of the spirometry results from the BOLD study provided this opportunity.

Methods

The design of BOLD is described in detail elsewhere[4] and only summarized here. A list of all participating entities in the BOLD Collaborative Research Group is included as an Addendum. Participating sites were expected to recruit a population-based sample of at least 600 non-institutionalized adults (300 women and 300 men), ages 40 and older, living in a well-defined administrative area (the "target population") whose total population exceeded 150,000. In this paper we report data from the first 14 BOLD sites. A more detailed description of these sites appears elsewhere.[5]

For this paper, data are limited to individuals with questionnaire data and both pre- and post-bronchodilator spirometry. The BOLD questionnaires included information on respiratory symptoms, risk factors for COPD, co-morbidities, and respiratory diagnoses, and were administered in face-to-face interviews by trained and certified staff. Each site obtained approval from their local ethical committee and written informed consent from each participant.

The same model of spirometer was purchased for each site. The spirometer model (ndd EasyOne Diagnostic model 2001, Zurich, Switzerland) was chosen to minimize the risk of cross-contamination, for portability, to provide automated quality checks and messages, and to store the results for transfer to a personal computer database. Spirometer calibration checks were done using a 3.00 L calibration syringe at a single speed every day of testing and once a month at 3 speeds (to check linearity), per ATS/ERS guidelines.[1] This spirometer has been demonstrated to remain accurate for prolonged periods of time.[6] Spirometry was performed before and 15 min after the administration (using a spacer) of 200 μg of albuterol/salbutamol.

At least one team member from each site was centrally trained for two days at the beginning of the program by a pulmonary specialist with considerable experience in spirometry testing. These technologists were qualified clinical technologists (at least 3 years training), nurses, or fieldworkers with no prior experience with spirometry. Spirometry was performed either in the homes of participants or at a local healthcare center. The participants were vigorously coached by the technicians to perform up to 8 FVC maneuvers until a quality grade of A or B was displayed on the spirometer. (See the Appendix for EasyOne quality grade details.) Results from the best 3 maneuvers (highest sum of FEV_1 plus FVC) were stored by the spirometer. Adequate quality was considered a grade of A, B, or C.

The quality of all test sessions was reviewed by the BOLD Pulmonary Function Reading Center (RJ). Quality reports were regularly sent to the BOLD clinical centers, as done in the Lung Health Study.[7] If the overall quality for the most recent 10 tests was considered sub-optimal, the site principal investigator was required to provide remedial spirometry training for the technologist(s) who were performing inadequately. It was recommended that the technician not perform additional testing for the study until retraining and recertification was completed.

To assess the overall quality of performance by our subjects and technicians for this paper, descriptive statistics were calculated for the maneuver acceptability variables BEV and PEFT from the maneuver with the highest FEV_1; EOTV and FET from the maneuver with the highest FVC; and for FVC and FEV_1 repeatability (dFVC and $dFEV_1$, highest minus second highest).[8]

To identify significant influences on performance, a logistic regression analysis was performed with adequate quality as the dependent variable. The initial regressions included age (in ten year increments), sex, smoking status (current or former smoker versus never smoker), education level (9 or more years of school), dyspnea (MRC grade 2 or higher), pre-BD FEV_1 (%predicted), BD responsiveness (% change from baseline), and study site. A p-value <0.02 was considered significant.

Airway obstruction was defined as FEV_1/FVC below the fifth percentile lower limit of the normal range (LLN)[9] and FEV_1 below 65% predicted, using NHANES III reference equations for Caucasians[10] (regardless of reported race or ethnicity). Spirometric restriction was defined as FVC below the LLN with FEV_1/FVC above the LLN.

Results

About half of the study participants were men; one-fourth were current smokers; one-third were former smokers;

Table 1 Characteristics and spirometry results from the 9893 participants (including those with poor quality spirometry tests): mean, 5th, and 95th percentiles.

	Mean	5th	95th
Age, yrs	56.5	41	78
Height, cm	166	150	184
Weight, Kg	75.1	50	105
BMI	27.1	19.8	36.7

	Pre-BD			Post-BD		
	Mean	5th	95th	Mean	5th	95th
FEV_1, L	2.66	1.27	4.19	2.74	1.37	4.30
FEV_1, % pred	88.6%	54.6	116.5	91.4%	59.2	18.9
FVC, L	3.60	1.93	5.56	3.60	1.97	5.54
FVC, % pred	92.8%	64.0	119.8	92.9%	65.8	119.9
FEV_1/FVC	0.74	0.57	0.85	0.76	0.59	0.87

one-fourth were obese (BMI >30); 7.1% reported current asthma; 11% reported MRC grade 2–5 dyspnea; 4.8% had airway obstruction; 20.4% had spirometric restriction, and the mean BD response was a 3.7% higher FEV_1 (SD 7.1%). Table 1 presents additional characteristics of the participants and descriptive statistics for the key spirometry variables. Overall, 96% of pre-bronchodilator test sessions and 97% of post-bronchodilator test sessions had adequate quality (EasyOne quality grades A, B, or C), although this varied amongst BOLD sites from a low of 86% to a high of 100% (Table 2).

The highest FEV_1 and the highest FVC are used to obtain the FEV_1/FVC which is used to determine airway obstruction, so the quality of these maneuvers is important. The accuracy of the FEV_1 in this ratio depends on a maximally deep breath and then a small BEV and PEFT. Table 3 shows that the BEV was below 154–158 mL in 90% of the maneuvers with the highest FEV_1; and the PEFT was less than 110 ms in 90% of these maneuvers. The accuracy of the FVC depends on a maximally deep breath and then a long maneuver with a small end-of-test volume, indicating a flat volume-time plateau. The FET was more than 6.4 s in 90% of the maneuvers with the highest FVC; and the EOTV was smaller than 26–30 mL in 90% of these maneuvers.

Table 2 Success rates for obtaining adequate quality (ndd EasyOne grade A, B, or C) by BOLD study site.

Site	Tests	Pre-BD	Post-BD
Sydney, Australia	585	88.3%	88.6%
Salzburg, Austria	1347	98.1%	97.3%
Vancouver, Canada	856	99.0%	98.3%
Guangzhou, China	590	97.7%	97.1%
Hannover, Germany	711	99.4%	98.3%
Reykjavik, Iceland	759	100%	100%
Lexington, Kentucky	559	86.4%	94.1%
Bergen, Norway	707	90.2%	93.6%
Manila, Philippines	918	99.8%	99.1%
Krakow, Poland	601	92.0%	95.5%
Cape Town, South Africa	896	95.6%	96.5%
Uppsala, Sweden	587	94.0%	94.3%
Adana, Turkey	864	98.6%	98.0%
London, U.K.	691	99.6%	97.8%

Unless the technologist is watching the subject during spirometry, the depth of the inhalation which preceded the forced exhalation can only be estimated by the repeatability of the FEV_1 and FVC (the highest value minus the second highest value). In 90% of the test sessions, the FEV_1s matched within 129–138 mL and the FVCs matched within 149–163 mL (pre- and post-BD respectively) (Table 3).

The ATS/ERS 2005 goals for spirometry quality for adults were set so that 90% of the patients seen in the pulmonary function laboratory of a large clinic could meet each of them.[11] Table 4 shows that each of the maneuver acceptability goals were also met in the single best maneuver by about 90% or more of the participants of the BOLD study. As expected, the study participants were slightly more successful during post-BD test sessions when compared to pre-BD test sessions (when they were test naïve). The ATS/ERS FEV_1 and FVC repeatability goal is <150 mL. BOLD participants met these goals in 92.4% and 87.4% of pre-BD test sessions ($dFEV_1$ and dFVC, respectively) and 93.5 and 90.3% of post-BD test sessions ($dFEV_1$ and dFVC respectively).

For the purpose of detecting "slow starts," addition of the PEFT threshold of >120 ms added very little to use of the traditional BEV threshold of >150 mL (Fig. 1). The BEV quality check was significantly more likely ($p < 0.001$) than PEFT to flag maneuvers as having a slow start (11–12% versus 7–8% failure rate, respectively). Only 5–6% of maneuvers with an acceptable BEV had an unacceptable PEFT. Maneuvers with a slow start (PEFT >120 ms) were significantly more likely ($p < 0.001$) to be stopped before 6 s (perhaps because the technologist quickly recognized the body language of the slow start and stopped the maneuver prematurely to re-instruct the participant).

Short maneuvers (FET < 6.0 s) in these adults were significantly more likely ($p < 0.001$) to lack have a flat volume-time plateau, as measured by an EOTV <40 mL (Fig. 2). About 39% of pre-BD maneuvers which lasted less than 4 s (and 15% of those with FET between 4 and 6 s) had a high EOTV, compared to just 3% for maneuvers lasting at least 6 s. Comparable post-BD figures were 23%, 9%, and 2.4%. In participants with clinically important airway obstruction post-BD, shorter maneuvers within a test session were significantly ($p < 0.001$) associated with smaller FVCs (Fig. 3).

Table 3 Results of quality checks from 9893 test sessions.

	Pre-BD				Post-BD			
	Mean	5th	90th	95th	Mean	5th	90th	95th
BEV, mL	99	36	158	185	98	39	154	178
PEFT, ms	85	60	110	120	84	60	110	120
FET, sec	10.1	5.3	6.5[a]		9.6	5.4	6.4[a]	
EOTV, mL	16	5	30	38	15	5	26	34
$dFEV_1$, mL	65	4	138	172	61	4	129	164
dFVC, mL	80	6	163	194	74	4	149	184

BEV and PEFT are from the maneuver with the highest FEV_1.
FET and EOTV are from the maneuver with the highest FVC.
[a] A short FET is not desirable, so the 10th percentile for FET is given instead of the 90th percentile.

Independent predictors of success in meeting quality goals for post-BD test sessions (EasyOne quality grade A, B, or C) included female sex, younger age, higher education, less dyspnea, a lower pre-BD FEV_1, less BD responsiveness, and study site (Table 5). However, these factors explained only about 10% of the overall variability in quality. Smoking status, MRC grade 1 dyspnea, and asthma were not significant independent predictors of spirometry quality. We did observe a significant age—sex interaction, but we removed this term from the final model to make the coefficients easier to understand. The odds ratios for acceptable quality from each of the other 13 sites (when compared to the Sydney, Australia site) ranged from 2.1 to 13.3 with all *p*-values $\leq$0.003.

Discussion

Despite a diversity of settings, languages, and ethnicities, these population-based samples of adults and the technologists who tested them achieved as high success rates in meeting spirometry quality goals as did patients seen in the PFT lab of a major referral medical center in the US (Mayo Clinic in Minnesota).[2,12] The thresholds specified by ATS/ERS 2005 standards[1] were set near the 90th percentile, so that about 10% of patients (both children and adults) fail to meet each criterion when tested by an experienced technician using a diagnostic quality spirometry system.[2]

Some *individual* BOLD sites were more likely to produce FVC maneuvers with short exhalation times (data not shown), which likely underestimated the prevalence of airway obstruction at that site. Submaximal inhalations cause under-estimates of the FVC and FEV_1. Poor blast efforts can cause under-estimates of the FEV_1,[13] while short exhalation times cause under-estimates of the FVC and the FEV_1/FVC.

The spirometry maneuver may be divided into 3 steps (or phases), each of which requires a different type of effort: 1) "take a deep breath" (maximal inhalation), 2) "blast out your air" (maximal exhalation effort), and 3) "keep blowing until all your air is gone" (prolonged exhalation). Poor effort may occur during any (or all) of these steps, and is usually due to sub-optimal interaction between the technologist and the subject. A submaximal inhalation falsely reduces all of the results (except for the ratios). A submaximal blast during the second phase reduces the measured PEF, variably affects the FEV_1, and may increase the FVC.[14] A premature termination of the exhalation falsely reduces the FVC (and the FEV6, if it occurs before 6 s), and is detected by a high end-of-test volume (EOTV).

Table 4 "Best maneuver" acceptability rates (per ATS/ERS 2005).

	pre-BD	post-BD
BEV <150 mL	88.1%	89.1%
PEFT <120 ms	91.8%	93.2%
FET >6sec	92.4%	92.6%
EOTV <40 mL	95.6%	96.9%

BEV and PEFT are from the maneuver with the highest FEV_1.
FET and EOTV are from the maneuver with the highest FVC.

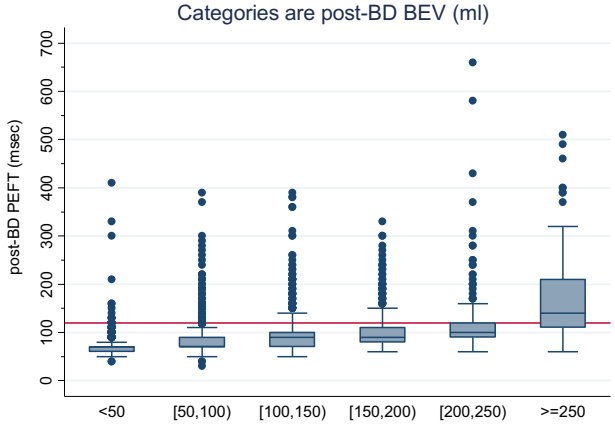

Figure 1 Box and whisker plots show the relationship between back-extrapolated volume (BEV in mL) and time to peak flow (PEFT in msec) for detecting slow starts in post-BD test sessions. BEV was categorized in 50 mL increments. A BEV >150 mL or a PEFT >120 ms indicates an unacceptably slow start. The bottom and top of each box indicate the 25th and 75th percentiles, while the bottom and top whiskers indicate the 5th and 95th percentiles. Only 19% of the variation in BEV was explained by the variation in PEFT.

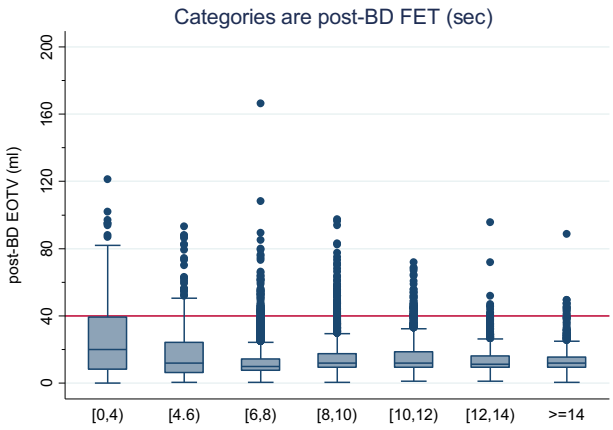

Figure 2 Box and whisker plots show the relationship between the end-of-test volume (EOTV in mL) and the forced expiratory time (FET) for detecting premature terminations of FVC maneuvers. ATS/ERS acceptability thresholds are <45 mL for EOTV and >6.0 s for FET. The bottom and top of each box indicate the 25th and 75th percentiles, while the bottom and top whiskers indicate the 5th and 95th percentiles. Note that after 6 s, fewer than 5% of maneuvers had an unacceptable EOTV (corresponding to a flat plateau on the volume-time curve).

Objective quality checks are designed to detect all of the above faults, and thereby to identify any poorly performed maneuver or test session which could result in false positive or false negative diagnoses in the clinical setting, or increased measurement noise/bias in epidemiologic and intervention studies. Poor inhalation effort is common, but is not objectively evident in any single spirometric record.

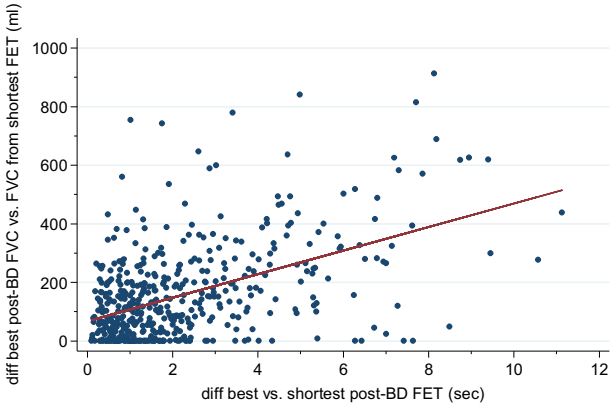

Figure 3 The relationship between shorter forced expiratory times and smaller FVCs within post-BD test sessions in the subset of 559 adults with clinically important airway obstruction (FEV_1/FVC below the lower limit of the normal range and FEV_1 below 65% predicted). Within each test session, the maneuver with the highest FVC was considered the best FVCb, which was then compared to the maneuver with the shortest FET (FVCs). dFVC (in milliliters on the vertical axis) = FVCb minus FVCs. About 37% of the differences in FVC were explained by shorter FETs. Similar relationships were seen pre-BD and in all study participants (but with lower R-squared values, data not shown).

Thus poor inhalation effort can be detected only in terms of poorly reproducible FVC and FEV_1 across multiple maneuvers. Submaximal blast and premature termination can, however, be identified objectively from the recording of any single blow.

The second phase of the spirometry maneuver is to BLAST out the air as quickly as possible, thereby achieving a "sharp" (high) peak flow during the first tenth of a second and a high average flow during the first second of the maneuver (FEV_1). A hesitating start creates a high back-extrapolated volume (BEV), causing an error in the measured FEV_1, so the ATS guidelines consider maneuvers with a high BEV to be unacceptable. A long time to reach peak flow (PEFT) indicates a relatively slow start, or lack of a maximal effort to blast out the air.

The ATS/ERS goal for a rapid start-of-test (BEV < 0.15 L, whichever is greater) was met in more than 90% of tests done. The software version of EasyOne used for this project (version 2.10) had a low pass filter designed to remove high frequency flow "noise" above 10 Hz. This filter reduced the PEFT and PEF (FEFmax) somewhat for very sharp blast efforts.

End-of-test maneuver acceptability criteria are designed to detect maneuvers which "quit too soon" resulting in an under-estimation of the true FVC. The ATS/ERS 2005 recommendations require FET >6 s for adults, and an "obvious plateau" in the volume-time curve. About 90% of the post-BD maneuvers with the highest FVC in our study achieved an EOTV of less than 29 mL.

The correlates of good quality spirometry have also been reported for other studies of adults.[2,3,12,15,16] The individual technologists performing the tests are the most common source of variability in quality. In the BOLD study, these technologist differences are represented by the study sites, since we did not ask technologists within each site to identify themselves for each test. Participant characteristics, such as female sex, younger age (within adulthood), and higher education which independently predict success in meeting quality goals are consistent between these studies and the BOLD study. Those with respiratory symptoms, airway obstruction, or bronchodilator responsiveness were slightly less likely to meet quality goals, especially for a flat volume-time curve (small EOTV). The lack of an independent association with a history of asthma with poor quality suggests that maneuver-induced bronchospasm is rare in adults with asthma.

This study adds unique information when compared to previous studies of spirometry quality.[17–22] Many countries were included in the BOLD study, making the results broadly generalizable, while other studies were performed in only one city or one country. This study used a modern flow-sensing spirometer with automated quality checks and messages (in contrast to some previous studies which used volume spirometers which make it easier to meet EOTV goals, since the exhaled air is cooling and contracting inside the spirometer during the final seconds of each maneuver). The most recent ATS/ERS goals for spirometry quality[1] were used by the BOLD study; and both pre- and post-BD test results were analyzed, making the results more applicable for COPD case-finding, which uses post-BD spirometry results.

Our results confirm and expand the results of previous studies of spirometry quality. A large study from Bergen Norway included only men ages 30–46 (younger than the

Table 5 Summary of logistic regression model[a] for predicting adequate post-bronchodilator spirometry.

Variable	Odds Ratio (multivariate)	95% CI	*p*-value	Odds Ratio (univariate)	95% CI	*p*-value
Female gender	1.89	(1.49, 2.40)	<0.001	1.61	(1.29, 2.01)	<0.001
Age (increase of 10 years)	0.83	(0.75, 0.92)	<0.0001	0.70	(0.64, 0.77)	<0.001
Current vs. never smoking	1.20	(0.87, 1.65)	0.26	1.27	(0.96, 1.70)	0.095
Ex vs. never smoking	1.29	(0.99, 1.69)	0.058	1.05	(0.82, 1.35)	0.68
MRC dyspnea level 1 vs. level 0	0.99	(0.68, 1.44)	0.94	1.01	(0.71, 1.45)	0.95
MRC dyspnea levels 2—3 vs. level 0	0.55	(0.35, 0.86)	0.008	0.67	(0.44, 1.00)	0.052
MRC dyspnea levels 4—5 vs. level 0	0.45	(0.28, 0.72)	0.001	0.50	(0.34, 0.76)	0.001
Unable to walk[b] vs. level 0	0.60	(0.42, 0.85)	0.005	0.50	(0.37, 0.68)	<0.001
Education grade 9—12 vs. 0—8	1.54	(1.14, 2.07)	0.004	1.28	(0.99, 1.64)	0.055
Education grade 13 + vs. 0—8	1.66	(1.16, 2.37)	0.005	1.67	(1.26, 2.22)	<0.001
Bronchodilator response (increase of 1% point)	0.96	(0.94, 0.97)	<0.0001	0.97	(0.95, 0.98)	<0.001
Pre-BD FEV_1, % (increase of 10% points)	0.88	(0.82, 0.94)	<0.0001	0.97	(0.92, 1.03)	0.34

[a] indicators of site ($p < 0.001$) were also included in the model though not shown in table.
[b] unable to walk for reasons other than lung disease.

age for which COPD case-finding is indicated) and only pre-BD spirometry.[18] About 90% of the almost 30,000 men in that study met the 1993 European goals for FEV_1 and FVC repeatability (<5% or <100 mL), which is the same rate as for BOLD participants meeting current repeatability goals. A more recent study from Bergen, Norway performed both pre- and post-BD spirometry for men and women around ages 47 and 72.[22] As in the BOLD study (which included 707 men and women from Bergen, Norway), male gender, older age, and dyspnea were independent predictors of poor quality post-BD spirometry. As in our study, they found slightly better quality post-BD when compared to pre-BD. They found that obesity and cognitive impairment (in the older study participants) were also associated with poorer quality, but we did not measure these factors. Our protocol suggested stopping after 8 maneuvers, but one-third of their participants were coached to perform more than 8 maneuvers, which was successful for obtaining good quality 98% of the time.

A 1992 study of spirometry quality in young adults (ages 20—45) working in Montreal offices found that 89% met the goal of FEV_1 repeatability within 100 mL.[19] Currently smoking women, men with a history of asthma or eczema, and men or women with positive methacholine challenge results were more likely to have poorer FEV_1 repeatability. In the six U.S. cities study of 8522 white adults, 91% met the goal of FEV_1 repeatability within 5% or 100 mL.[20] Dyspnea and a history of asthma were independent predictors of poorer quality. While we did not perform inhalation challenge tests, we found that BD responsiveness (which is associated with asthma and bronchial responsiveness) was an independent predictor of poorer quality spirometry.

Limitations of this study

The EasyOne spirometer did not label maneuvers with FET <6sec or EOTV >40 mL as unacceptable. Our criterion for acceptable quality (A, B, or C quality grade from the EasyOne spirometer) included test sessions which did not meet ATS/ERS 2005 goals for quality, which require FEV_1 and FVC repeatability within 150 mL (equivalent to an EasyOne quality grade A). Our results may not apply to subjects under age 40. Measurements of PEFT and EOTV may differ when using other models of spirometers. Some of the relationships will vary when only patients with severe obstruction or restriction are studied. Quality was probably enhanced by careful training before the field work began, followed by monthly feedback from the Reading Center. The Reading Center reviews found maneuver errors (such as zero flow errors and the lack of a volume-time plateau) which were reported back to the study sites but not reflected in the ndd quality grades which were analyzed for this manuscript.

Summary

We found that 90% or more of our participants were successfully coached to perform pre-BD forced expiratory spirometric maneuvers which met acceptability criteria for BEV, EOTV, FET; and within-test session maneuver repeatability for FEV_1 and FVC. Participant characteristics can influence performance; but their overall effect is small with well-trained technicians who have experience.

Appendix: EasyOne diagnostic spirometer model maneuver quality checks (firmware version 2.10)

The message "Don't hesitate" was displayed when the BEV was higher than 150 mL (or 5%, whichever was greater). If the time to peak flow (PEFT) was <120 ms, the message was "Blast out faster." The maneuver was marked as unacceptable by the spirometer if either of these thresholds were exceeded.

If the end-of-test volume (EOTV) was above 45 mL during the final 2 s, or when the BEV was >100 mL during the final 0.5 s when the forced expiratory time (FET) was less than 6 s for an adult, the message "Blow out longer" was displayed. These criteria were not used by the spirometer to determine acceptability.

Test Session Quality Grades: The quality of each spirometry test session was graded as follows (displayed after each maneuver and printed on the report):

A = 3 + acceptable maneuvers, AND FEV_1 and FVC match within 150 mL.

B = 3 + acceptable maneuvers, AND FEV_1 and FVC match within 200 mL.

C = 2 + acceptable maneuvers, AND FEV_1 and FVC match within 250 mL.

D = Only 1 acceptable maneuver, OR the FEV_1 or the FVC from the best 2 acceptable maneuvers do not match within 250 mL.

F = No acceptable maneuvers.

The message "Session Complete! Good Job" was displayed with a grade of A or B after 3 or 4 maneuvers, or C or better after 5 or more maneuvers.

Footnote: Some of these quality criteria were changed for EasyOne spirometers manufactured after September 2007 (with firmware versions 2.17 and higher).

Abbreviations

ATS	American Thoracic Society
BEV	back-extrapolated volume
BMI	body mass index
$dFEV_1$	the difference between the highest and second highest FEV_1 within a spirometry test session
EOTV	end-of-test volume (mL exhaled during the final 2.0 s)
FET	forced expiratory time
PEF	peak expiratory flow (as determined by spirometry)
PEFT	peak expiratory flow time (the time in milliseconds from back-extrapolated time zero until the peak flow occurs)
QA	Quality assurance

Addendum

The Burden of Obstructive Lung Disease (BOLD) Collaborative Research Group.

Executive Committee — A. Sonia Buist, chair (Oregon Health and Sciences University, Portland, OR, USA); Peter Burney (National Heart and Lung Institute, Imperial College, London, UK); Todd Lee (Northwestern University, Chicago, IL, USA); David M. Mannino (University of Kentucky, Lexington, KY, USA); Mary Ann McBurnie (Kaiser Permanente Center for Health Research, Portland, OR, USA); Ana MB Menezes (Federal University of Pelotas, Brazil); Sean Sullivan (University of Washington, Seattle, WA, USA); Jørgen Vestbo (Hvidovre University Hospital, Hvidovre, Denmark); William M. Vollmer (Kaiser Permanente Center for Health Research, Portland, OR, USA); Kevin B. Weiss (Northwestern University, Chicago, IL, USA).

International Advisory Board — Amund Gulsvik, chair (University of Bergen, Bergen, Norway); Josep M Antó (Center for Research in Environmental Epidemiology [CREAL, IMIM], Institut Municipal d'Investigació Mèdica (IMIM), Spain); Robert Crapo (Latter Day Saints Hospital, Salt Lake City, UT, USA); Guy Marks (Woolcock Institute of Medical Research, Sydney, Australia); Rogelio Perez-Padilla (Instituto Nacional de Enfermedades Respiratorias, Mexico City, Mexico); Gregory R. Wagner (CDC/NIOSH, Washington, DC, USA).

Operations Center — William M. Vollmer (PI), Michael Allison, Paul Cheek, Linda Figurski, E. Ann Frazier, Suzanne Gillespie, Chris Kelleher, Terry Kimes, Nidhi Kochar, Mary Ann McBurnie, Gayle Thomas-Monk, Esma Vance (Kaiser Permanente Center for Health Research, Portland, OR, USA); A. Sonia Buist (Oregon Health and Sciences University, Portland, OR, USA); Virgina Lesser (Oregon State University).

Economics Core — Todd Lee and Kevin B. Weiss (Northwestern University, Chicago, IL, USA); Sean D. Sullivan, (University of Washington, Seattle, WA, USA);

Pulmonary Function Reading Center — Bob Crapo, Robert Jenson (Latter Day Saints Hospital, Salt Lake City, UT, USA).

Field Centers — NanShan Zhong (PI), Shengming Liu, Jiachun Lu, Pixin Ran, Dali Wang, Jingping Zheng, Yumin Zhou (***Guangzhou Institute of Respiratory Diseases, Guangzhou Medical College, Guangzhou, China***); Ali Kocabaş (PI), Attila Hancioglu, Ismail Hanta, Sedat Kuleci, Ahmet Sinan Turkyilmaz, Sema Umut, Turgay Unalan (***Cukurova University School of Medicine, Department of Chest Diseases, Adana, Turkey***); Michael Studnicka (PI), Bernd Lamprecht, Lea Schirnhofer (***Paracelsus Medical University, Department of Pulmonary Medicine, Salzburg Austria***); Eric Bateman (PI), Anamika Jithoo (PI), Desiree Adams, Edward Barnes, Jasper Freeman, Anton Hayes, Sipho Hlengwa, Christine Johannisen, Mariana Koopman, Innocentia Louw, Ina Ludick, Alta Olckers, Johanna Ryck, Janita Storbeck, (***University of Cape Town Lung Institute, Cape Town, South Africa***); Thorarinn Gislason (PI), Bryndis Benedikdtsdottir, Kristin Börundsdottir, Lovisa Gudmundsdottir, Sigrun Gudmundsdottir, Gunnar Gundmundsson, (***Landspitali University Hospital, Dept. of Allergy, Respiratory Medicine and Sleep, Reykjavik, Iceland***); Ewa Nizankowska-Mogilnicka (PI), Jakub Frey, Rafal Harat, Filip Mejza, Pawel Nastalek, Andrzej Pajak, Wojciech Skucha, Andrzej Szczeklik,Magda Twardowska, (***Division of Pulmonary Diseases, Department of Medicine, Jagiellonian University School of Medicine, Cracow, Poland***); Tobias Welte (PI), Isabelle Bodemann, Henning Geldmacher, Alexandra Schweda-Linow (***Hannover Medical School, Hannover, Germany***); Amund Gulsvik (PI), Tina Endresen, Lene Svendsen (***Department of Thoracic Medicine, Institute of Medicine, University of Bergen, Bergen, Norway***); Wan C. Tan (PI), Wen Wang (***iCapture*** Center for Cardiovascular and Pulmonary Research, University of British Columbia, Vancouver, BC, Canada); David M. Mannino (PI), John Cain, Rebecca Copeland, Dana Hazen, Jennifer Methvin, (***University of Kentucky, Lexington, Kentucky, USA***); Renato B. Dantes (PI), Lourdes

Amarillo, Lakan U. Berratio, Lenora C. Fernandez, Norberto A. Francisco, Gerard S. Garcia, Teresita S. de Guia, Luisito F. Idolor, Sullian S. Naval, Thessa Reyes, Camilo C. Roa, Jr., Ma. Flordeliza Sanchez, Leander P. Simpao (***Philippine College of Chest Physicians, Manila, Philippines***); Christine Jenkins (PI), Guy Marks (PI), Tessa Bird, Paola Espinel, Kate Hardaker, Brett Toelle (***Woolcock Institute of Medical Research, Sydney, Australia***).

Acknowledgements & role of study sponsors

The BOLD initiative has been funded in part by unrestricted educational grants to the Operations Center (www.boldcopd.org) from ALTANA, Aventis, AstraZeneca, Boehringer-Ingelheim, Chiesi, GlaxoSmithKline, Merck, Novartis, Pfizer, Schering-Plough, Sepracor, and University of Kentucky. Additional local support for BOLD clinical sites was provided by: Boehringer Ingelheim China. (***GuangZhou, China***); Turkish Thoracic Society, Boehringer-Ingelheim, and Pfizer (***Adana, Turkey***); Altana, Astra-Zeneca, Boehringer-Ingelheim, GlaxoSmithKline, Merck Sharpe & Dohme, Novartis, Salzburger Gebietskrankenkasse and Salzburg Local Government (***Salzburg, Austria***); Research for International Tobacco Control, the International Development Research Center, the South African Medical Research Council, the South African Thoracic Society GlaxoSmithKline Pulmonary Research Fellowship, and the University of Cape Town Lung Institute (***Cape Town, South Africa***); and Landspítali-University Hospital-Scientific Fund, GlaxoSmithKline Iceland, and AstraZeneca Iceland (***Reykjavik, Iceland***); GlaxoSmithKline Pharmaceuticals, Polpharma, Ivax Pharma Poland, AstraZeneca Pharma Poland, ZF Altana Pharma, Pliva Kraków, Adamed, Novartis Poland, Linde Gaz Polska, Lek Polska, Tarchomińskie Zakłady Farmaceutyczne Polfa, Starostwo Proszowice, Skanska, Zasada, Agencja Mienia Wojskowego w Krakowie, Telekomunikacja Polska, Biernacki, Biogran, Amplus Bucki, Skrzydlewski, Sotwin, and Agroplon (***Cracow, Poland***); Boehringer-Ingelheim, and Pfizer Germany (***Hannover, Germany***); the Norwegian Ministry of Health's Foundation for Clinical Research, and Haukeland University Hospital's Medical Research Foundation for Thoracic Medicine (***Bergen, Norway***); AstraZeneca, Boehringer-Ingelheim, Pfizer, and GlaxoSmithKline (***Vancouver, Canada***); Marty Driesler Cancer Project (***Lexington, Kentucky, USA***); Altana, Boehringer Ingelheim (Phil), GlaxoSmithKline, Pfizer, Philippine College of Chest Physicians, Philippine College of Physicians, and United Laboratories (Phil) (***Manila, Philippines***); Air Liquide Healthcare P/L, AstraZeneca P/L, Boehringer Ingelheim P/L, GlaxoSmithKline Australia P/L, Pfizer Australia P/L (***Sydney, Australia***).

The **study sponsors** had no involvement in the study design, data collection, analysis and interpretation of data, in the writing of the manuscript, and in the decision to submit the manuscript for publication.

Conflict of interest statement

The authors declare that they have no conflict of interest.

References

1. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Crapo R, et al. ATS/ERS task force. Standardisation of spirometry. *Eur Respir J* 2005;**26**(2):319—38.
2. Enright PL, Beck KC, Sherrill DL. Repeatability of spirometry in 18,000 adult patients. *Am J Respir Crit Care Med* 2004;**169**:235—8.
3. Enright PL, Skloot GS, Cox-Ganser JM, Udasin IG, Herbert R. Quality of spirometry performed by 13,599 participants in the world trade center worker and volunteer medical screening program. *Respir Care* 2010;**55**(3):303—9.
4. Buist AS, Vollmer WM, Sullivan SD, Weiss KB, Lee TA, Menezes AM, et al. The Burden of obstructive lung disease initiative (BOLD): rationale and design. *COPD* 2005;**2**(2):277—83.
5. Buist AS, McBurnie MA, Vollmer WM, Gillespie S, Burney P, Mannino DM, et al. International variation in the prevalence of COPD (the BOLD Study): a population-based prevalence study. *Lancet* 2007;**370**(9589):741—50.
6. Pérez-Padilla R, Vázquez-García JC, Márquez MN, Jardim JR, Pertuzé J, Lisboa C, et al. for the Latin American COPD prevalence study (PLATINO) team. The long-term stability of portable spirometers used in a multinational study of the prevalence of chronic obstructive pulmonary disease. *Respir Care* 2006;**51**(10):1167—71.
7. Enright PL, Johnson LR, Connett JE, Voelker H, Buist AS. Spirometry in the lung health study: 1. methods and quality control. *Am Rev Respir Dis* 1991;**143**:1215—23.
8. Enright PL, Connett JE, Kanner RE, Johnson LR, Lee WW. Spirometry in the lung health study: II. Determinants of short-term intra-individual variability. *Am J Respir Crit Care Med* 1995;**151**(2 Pt 1):406—11.
9. Vollmer WM, Gíslason T, Burney P, Enright PL, Gulsvik A, Kocabas A, et al. Comparison of spirometry criteria for the diagnosis of COPD: results from the BOLD study. *Eur Respir J* 2009;**34**:588—97.
10. Hankinson J, Odencrantz J, Fedan K. Spirometric reference values from a sample of the general U.S. population. [NHANES III]. *Am J Respir Crit Care Med* 1999;**159**:179—87.
11. Hankinson JL, Bang KM. Acceptability and reproducibility criteria of the ATS for spirometry as observed in a sample of the general population. *Am Rev Respir Dis* 1991;**143**:516—21.
12. Herpel LB, Kanner RE, Lee SM, Fessler HE, Sciurba FC, Connett JE, et al. Variability of spirometry in chronic obstructive pulmonary disease: results from two clinical trials. *Am J Respir Crit Care Med* 2006;**173**(10):1106—13.
13. Hegewald MJ, Lefor MJ, Jensen RL, Crapo RO, Kritchevsky SB, Haggerty CL, et al. Peak expiratory flow is not a quality indicator for spirometry: peak expiratory flow variability and FEV_1 are poorly correlated in an elderly population. *Chest* 2007; **131**(5):1494—9.
14. Krowka MJ, Enright PL, Rodarte JR, Hyatt RE. Effect of effort on measurement of forced expiratory volume in one second. *Am Rev Respir Dis* 1987;**136**:829—33.
15. Humerfelt S, Eide GE, Kvale G, Gulsvik A. Predictors of spirometric test failure: a comparison of the 1983 and 1993 acceptability criteria from the ECCS. *Occup Environ Med* 1995; **52**:547—53.
16. Kunzli N, Ackermann-Liebrich U, Keller R, Perruchoud AP, Schindler C. Variability of FVC and FEV_1 due to technician, team, device, and subject in an eight center study: three quality control studies in SAPALDIA. Swiss study on air pollution and lung disease in adults. *Eur Respir J* 1995;**8**:371—6.
17. Lorber DB, Kaltenborn W, Burrows B. Responses to isoproterenol in a general population sample. *Am Rev Respir Dis* 1978; **118**:855—61.
18. Humerfelt S, Eide GE, Kvåle G, Gulsvik A. Predictors of spirometric test failure: a comparison of the 1983 and 1993

acceptability criteria from the European community for coal and steel. *Occup Environ Med* 1995;**52**:547—53.
19. Ng'Ang'a LW, Ernst P, Jaakkola MS, Gerardi G, Hanley JH, Becklake MR. Spirometric lung function. Distribution and determinants of test failure in a young adult population. *Am Rev Respir Dis* 1992;**145**:48—52.
20. Eisen EA, Dockery DW, Speizer FE, Fay ME, Ferris Jr BG. The association between health status and the performance of excessively variable spirometry tests in a population-based study in six U.S. cities. *Am Rev Respir Dis* 1987;**136**:1371—6.
21. Kainu A, Lindquist A, Sarna S, Lundbäck B, Sovijärvi A. FEV_1 response to bronchodilation in an adult Urban population. *CHEST* 2008;**134**:387—93.
22. Lehmann S, Vollset SE, Nygaard HA, Gulsvik A. Factors determining performance of bronchodilator reversibility tests in middle-aged and elderly. *Respir Med* 2004;**98**:1071—9.